Heritability of attention problems in children II: longitudinal results from a study of twins age 3 to 12.

Rietveld, M.J.H.; Hudziak, J.; Bartels, M.; van Beijsterveldt, C.E.M.; Boomsma, D.I.

published in
Journal of Child Psychology and Psychiatry
2004

DOI (link to publisher)
10.1111/j.1469-7610.2004.00247.x

document version
Publisher's PDF, also known as Version of record

Journal of Child Psychology and Psychiatry 45:3 (2004), pp 577–588

Heritability of attention problems in children: longitudinal results from a study of twins, age 3 to 12

M.J.H. Rietveld,[1] J.J. Hudziak,[2] M. Bartels,[1] C.E.M. van Beijsterveldt,[1] and D.I. Boomsma[1]

[1]Department of Biological Psychology, Vrije Universiteit, Amsterdam, The Netherlands; [2]Department of Psychiatry and Medicine (Division of Human Genetics), Center for Children, Youth and Families, University of Vermont, USA

Background: Twin studies of childhood behavior problems support the conclusion that individual differences in impulsivity, hyperactivity, and inattention are largely due to genetic influences. Non-genetic variation is due to environmental influences that are unique to the individual, and possibly to rater contrast effects. In the present longitudinal twin study, we report on the size of genetic and environmental effects on individual differences in attention problems at ages 3, 7, 10 and 12 years. **Methods:** Mothers were asked to complete the CBCL for their twin offspring when the children were 3 (n = 11,938), 7 (n = 10,657), 10 (n = 6,192), and 12 years old (n = 3,124). We focus on the Overactivity (OA) scale in the Child Behavior Checklist (CBCL/2–3), and on the Attention Problem (AP) scale of the CBCL/4–18. The data were analyzed using longitudinal structural equation modeling. **Results:** Broad heritability of OA and AP is estimated at nearly 75%, at each age. A contrast effect was observed at age 3 only. The results revealed less stability of OA at age 3 to AP at age 7 (r = .40), compared to the stability from AP at age 7 and beyond (r = .70). Genetic effects explained between 76% and 92% of the covariance between OA and AP. **Conclusions:** OA and AP are highly heritable at all ages in both genders. The same set of genes appears to be expressed in boys and girls. The size of genetic and environmental contributions remains the same across the ages studied. Stability in OA and AP is accounted for by genetic influences. Children who do not display OA or AP at a given age are unlikely to develop these problems at a subsequent age. **Keywords:** Overactivity, attention problems, heritability, twin study, repeated measures. **Abbreviations:** OA = overactivity; AP = attention problems; CBCL = Child Behavior Checklist.

Attention Problems (AP) and Overactive behavior (OA) represent syndromes that can be measured by parental ratings on the Child Behavior Checklist (CBCL). OA was identified as a syndrome in factor analyses of responses of parents of Dutch pre-school children (Achenbach, 1992; Koot, Van den Oord, Verhulst, & Boomsma, 1997; Van den Oord, 1993). OA is predictive of, and developmentally related to, AP, a syndrome that emerged in factor analyses of responses of parents rating their children aged 6 to 18 years (Achenbach, 1991). In earlier work (Rietveld, Hudziak, Bartels, Van Beijsterveldt, & Boomsma, 2003a) we demonstrated that genetic influences on maternally rated OA in 3-year-olds explain 75% of the phenotypic variance. Genetic influences remained the most important source of variance on maternally rated AP at ages 7, 10 and 12 years. In this previous work cross-sectional analyses were carried out. Here, we present results of longitudinal analyses of these and additional data.

To date, several behavior genetic studies have incorporated the CBCL. Edelbrock, Rende, Plomin, and Thompson (1995) reported a heritability estimate of 66% for AP measured in twins aged 7 to 15 years. Similar estimates were reported by Schmitz, Saudino, Plomin, Fulker, and DeFries (1996) for 7-year-old children, by Zahn-Waxler, Schmitz, Fulker, Robinson, and Emde (1996) for 5-year-old children, and by Hudziak, Rudiger, Neale, Heath, and Todd (2000) for children in an age range from 8 to 12 years. The study by Gjone, Stevenson, and Sundet (1996) focused on two age ranges, 5 to 9 years and 12 to 15 years, and reported equal results over this period. Van den Oord, Verhulst, and Boomsma (1996) studied OA in 3-year-old Dutch twins, and found that more than half of the observed variation was explained by genetic influences. Their sample is included in the present study. Behavior genetic studies have also addressed the genetic and environmental influences on phenotypes other than OA and AP, but associated with ADHD (Kuntsi, Gayán, & Stevenson, 2000; Martin, Scourfield, & McGuffin, 2002; Thapar, Hervas, & McGuffin, 1995; Saudino, Cherny, & Plomin, 2000; Schmitz, Fulker, & Mrazek, 1995; Sherman, Iacono, & McGue, 1997). These studies have produced varying results, which may be due to the use of different assessment instruments (Nadder, Silberg, Rutter, Maes, & Eaves, 2001; Thapar, Harrington, Ross, & McGuffin, 2000). However, there is considerable consistency in the results. The larger studies including 1,000 or more twin pairs (Eaves et al., 1997; Nadder, Silberg, Eaves, Maes, & Meyer, 1998; Price et al., submitted; Rhee, Waldman, Hay, & Levy, 1999) reported heritability estimates between 60% and 85%. The environmental effects, which account for most of the remaining variance, are unique or unshared. The unique environment encompasses experiences that are specific to an individual and

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

make members of the same family different from one another. Environmental experiences, which are shared among family members, are negligible in explaining variation in overactivity, inattention, and impulsivity. Another consistent finding is the presence of a contrast effect in parental data on ADHD phenotypes. This effect shows up as a negative interaction between the twins' phenotypes (Carey, 1986; Eaves, 1976). The interaction appears to be due to bias, which originates in the rating of the parents, rather than a true social interaction between twins (Simonoff et al., 1998). When the contrast effect is present but not included in the analyses, heritability estimates are overestimated. In our previous report, contrast effects were detected at age 3, but the results were inconclusive at older ages. However, contrast effects are very difficult to distinguish from genetic dominance effects in studies of MZ and DZ twins only (Rietveld, Posthuma, Dolan, & Boomsma, 2003b).

In twin studies of AP and ADHD phenotypes, gender differences in heritability have received little attention, and the available results are largely inconclusive. One large study found gender effects in the magnitude of genetic and environmental influences on ADHD (Rhee et al., 1999), whereas two large twin studies reported the absence of gender effects for ADHD and related symptoms (Nadder et al., 1998; Thapar et al., 2000). In our cross-sectional study we found similar heritability estimates in boys and girls at ages 3, 7, 10, and 12 years.

From the literature on developmental child psychology (Biederman, 1998; Biederman, Mick, & Faraone, 2000; Ross & Ross, 1982; Rutter, Giller, & Hagell, 1998) it becomes clear that ADHD phenotypes follow a distinct pattern from infancy to early adolescence. Hyperactivity may be prominent during pre-school age and early school age, but it is less prevalent in older subjects. The inability to sustain attention is most profound during the school age years, but tends to decline in early adolescence, when antisocial behavior becomes more pronounced.

In order to study the pathways into and out of AP, Koot (1993) explored the relations of the Dutch OA syndrome in a community sample of 2- and 3-year-old children, and AP measured at follow-up two years later. Achenbach and Rescorla (2000) performed identical analyses for an adjusted AP scale in pre-school children and AP in school-aged children. The inter-occasion correlations of .40 and .47 reported in these studies suggest that stability is moderate. These correlations are lower than the .60 obtained by Verhulst and Van der Ende (1995) for AP itself over a period that covered 2 and 4 years.

With the exception of the studies by Schmitz et al. (submitted) and Van der Valk, Verhulst, Neale, and Boomsma (1998), longitudinal twin studies of AP and ADHD are lacking. Schmitz et al. (submitted) explored the stability and covariance of temperament and attention problems as measured with the CBCL at ages 7 and 12 years. The phenotypic correlation of maternally reported AP between these ages was .54. This stability in AP was explained mainly by genetic influences. These results are comparable to those reported in a study of adolescent adoptees, where phenotypic stability in AP during adolescence was estimated around .60 (Van der Valk et al., 1998). It was found that genetic effects and environmental effects contributed equally to the stability of attention problems during adolescence. As reported in the study of Schmitz et al. (submitted), environmental effects were unique to the individual.

The contributions of genetic and environmental effects to the (in)stability of OA and AP during childhood remain to be investigated. The aim of the present twin study is to evaluate the genetic and environmental effects on overactivity and attention problem behavior during childhood, in boys and girls. OA was measured in 3-year-old twins, and AP was measured in the same twins at ages 7, 10, and 12 years. This twin study is the first to cover a time period of this length. By using structural equation modeling of the longitudinal twin data, we address the following issues. Is phenotypic stability attributable to genetic influences, environmental influences, or a combination of both? Are the contributions of genetic and environmental factors stable across the entire age range (age 3 to age 12), or do they change during development? Are there gender differences in the genetic and environmental contributions to the inter-occasion correlations?

Materials and methods

Subjects and procedure

The Netherlands Twin Registry (NTR; Boomsma, 1998) was established in 1986 and is maintained by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam. For the present study, we assessed a sample of Dutch twin pairs, whose parents reported on their behavior at ages 3, 7, 10, and 12 years. Data from 3- and 7-year-old twins were collected from birth cohorts 1986 through 1993. The dataset of the 10-year-olds comprised data from birth cohorts 1986 through mid-1991. Data of 12-year-old twins were available from birth cohorts 1986 through mid-1989/1990. We collected 11,938 maternal reports for the 3-year-olds, 10,657 maternal reports for the 7-year-olds, 6,192 maternal reports for the 10-year-olds, and 3,124 maternal reports of the 12-year-olds.

Discriminant analysis of questionnaire items was used to classify the same-sex pairs into MZ and DZ twins (Rietveld et al., 2000). The number of participating twins by maternal report and their assigned zygosity are listed in Table 1.

The number of individuals with complete maternal data at all ages is 2,192. At each age, around 2% of the twins suffered from a disability or disease that

Table 1 Number of individual twins separately by zygosity status

	Age 3	Age 7	Age 10	Age 12
Unknown	282	11	1	2
MZM	1870	1771	1054	579
DZM	1854	1692	954	472
MZF	2146	2012	1249	638
DZF	1743	1639	933	466
DOS	3784	3289	1835	876
Total	11679	10414	6026	3033

Note: The number of individuals with unknown zygosity at age 3 is due to the nonparticipation of these families at later ages. At age 3, these families provided incomplete information on the zygosity questionnaire.

interfered severely with their daily functioning. These twins displayed more than twice as much problem behavior across the entire age range compared to healthy twins. In addition, the co-twins, although unaffected, displayed increased total problem scores. These twin pairs were dropped from the sample.

Parental occupation was assessed using a 5-point scale, ranging from manual labor to academic employment. The level of occupation of the parents of the twins is somewhat higher than the level in the general population (Centraal Bureau voor de Statistiek, 2002a). This may be due to the minimal educational level that is required to read and complete questionnaires, or it may reflect a difference between adults with children and adults in the population as a whole.

Measures

The Child Behavior Checklist (CBCL) is a standardized questionnaire for parents to rate the frequency and intensity of behavioral and emotional problems exhibited by their children in the past six months. Parents rate each item on a 3-point scale. The total problem score is obtained by summing the item scores. The questionnaire that is used for the 3-year-olds (CBCL/2-3; Achenbach, 1992) differs in item content from the questionnaire that is used for the 7-, 10-, and 12-year-olds (CBCL/4-18; Achenbach, 1991). Koot et al. (1997) report the Dutch syndrome scales of the CBCL/2-3, and comparability of these with the American syndrome scales. An overactive scale was identified in a Dutch dataset, which has no American equivalent. Syndrome scales of the CBCL/4-18 were composed according to the profile developed by Achenbach (1991; De Groot, Koot, & Verhulst, 1994). The content of the Dutch overactive scale and attention problem scale is given in Appendix I.

Statistical analyses

The statistical modeling program Mx (Neale, 1997) was used to calculate descriptive statistics of OA and AP, and to test for mean differences between genders and across age. In these tests, likelihood-ratio chi-square (χ^2) tests were used. These tests control for the dependency that exists between the scores from twins. Selective nonparticipation at ages 7, 10, and 12 was evaluated by application of the general linear model (GLM) with repeated measures in SPSS. In these analyses, the OA score of a twin was considered a repeated measure of his or her co-twin's OA score. Participation at follow-up ages (yes/no) was specified as the between-group factor. Selective nonparticipation may change the composition of the sample at ages 7, 10, and 12 compared to the original sample at age 3. A non-significant effect implies that the age-specific samples may be considered to be 'selected completely at random' in the sense of Little and Rubin (1987). If nonparticipation is related to the OA score in the beginning of the study, one may still assume that the sample is 'selected at random'. This assumption is tenable when the willingness of mothers to return the questionnaire depended on variables other than OA but in some way related to OA. If the sample is 'selected completely at random' or 'selected at random', correct maximum likelihood (ML) estimates are obtained if raw data likelihood estimation is used and all available data are included (Wothke, 2000). Here, raw data ML estimation was carried out using Mx. This method requires that the data are normally distributed. The data were square root transformed to render their distribution approximately normal.

Genetic modeling

Figure 1 depicts the baseline model at one time point used for the analyses. The circles represent the latent, unmeasured variables and the squares represent the observed and measured phenotypes.

Phenotypic variance is modeled as the sum of genetic and environmental variances. We distinguish two categories of genetic sources of variance: additive effects of alleles at a large number of loci (A), and non-additive genetic effects reflecting interaction effects between alleles of the same gene locus (dominance, D). The environmental source of variance of OA and AP is con-

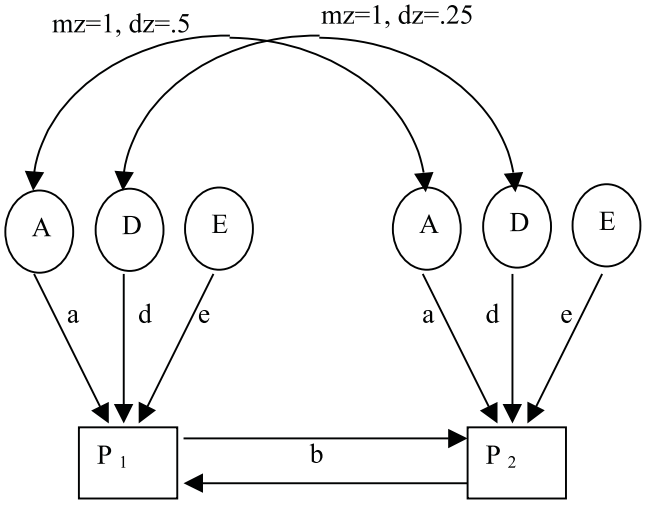

Figure 1 Path diagram of ADE model with contrast effect. P_1 and P_2 represent the observed phenotypes of twin 1 and twin 2, respectively. A, D, and E are additive genetic, genetic dominance, and unique environmental latent factors, respectively; a, d, and e are path-coefficients; b represents the negative contrast effect. Genetic correlations between twins are represented by double-headed arrows between the A and D factors.

sidered to be unique to the individual. These environmental influences (denoted E in Figure 1) contribute to differences between individuals within the same family. Estimates of these unique environmental contributions to the total variance include measurement error. The coefficients a, d, and e are factor loadings of P on the latent factors A, D and E. Given that the latent factors are constrained to have unit variance (Neale et al., 1992), the decomposition of the phenotypic variance (Vp) is:

$$Vp = a^2 + d^2 + e^2$$

The different degree of genetic relatedness of monozygotic (MZ) and dizygotic (DZ) twins is used to identify the contributions of the latent factors to the phenotypic variance of OA and AP (Plomin, DeFries, McClearn, & McGuffin, 2001). Similarity (covariances) of MZ twins may be due to additive genetic influences and genetic dominance, that is, $a^2 + d^2$. DZ twins share on average 50% of their segregating genes. The expectation for the DZ correlations between the genetic additive and dominance deviations are 1/2 and 1/4, respectively (Falconer & Mackay, 1996). Thus, the total DZ covariance equals $1/2\ a^2 + 1/4\ d^2$. Non-shared environmental influences (e^2) do not contribute to the covariance between individuals.

In this univariate model (Figure 1) path *b* is specified between the phenotypic scores of the twins. This path represents a reciprocal interaction between twins (Carey, 1986; Eaves, 1976; Neale & Stevenson, 1989; Simonoff et al., 1998). The literature on ADHD and associated phenotypes indicates that this interaction is negative (i.e., a contrast effect), and that it most likely represents a rater bias (Eaves et al., 2000; Hudziak, 2001; Kuntsi et al., 2000; Martin et al., 2002; Nadder et al., 1998; Price et al., submitted; Rietveld et al., 2003a; Thapar et al., 1995). As the parameter *b* is found to be negative, this suggests that parents stress the differences in behavior between their twins when rating their behavior.

If multiple assessments of the same twins are available, the covariance between measurements can also be modeled as a function of A, D and E (Boomsma & Molenaar, 1986; Martin & Eaves, 1977). When the same twins are measured repeatedly across time, one can evaluate how genes and environment contribute to the observed phenotypic stability during development. Stability is expressed as the correlation or covariance between the repeated measures. To accommodate the observed covariance between measures over time, the age-specific ADE univariate model was extended. Using a Cholesky or triangular decomposition model, the phenotypic covariance of OA and AP can be decomposed into additive genetic, genetic dominance, and unique environmental variance that is common across measurement occasions, and into additive genetic, genetic dominance, and unique environmental variance that is unique to each measurement occasion. The common variance is indicative of the extent to which additive genetic, genetic dominance, and unique environmental variance are shared across measurement occasions. The unique variance is indicative of the extent to which variance is age-specific, that is, not shared across ages. A contrast effect was specified at each age.

Model fitting procedures

Cross-sectional analyses of this and earlier datasets (Rietveld et al., 2003a) produced largely the same results. Longitudinal data analysis of measures collected at all four ages were done by a Cholesky decomposition. This model served as a reference model to judge the significance of genetic dominance and contrast effects.

Alternative models were evaluated on the basis of their plausibility and goodness-of-fit. Restrictions that are imposed in an alternative model are evaluated by minus twice the difference in log-likelihood between the restricted and the more general model. This difference is asymptotically chi-square distributed. If the χ^2 test is not significant, and the alternative model provides a plausible account of the data, the alternative model is considered tenable (Loehlin, 1992).

At all stages of model fitting the parameters were free to differ over gender. To evaluate whether boys and girls share a common set of genes, we tested whether opposite-sex twin correlations were of identical magnitude to same-sex dizygotic correlations. A discrepancy between these sets of correlations points at separate genetic sources of (co)variance for boys and girls, that is, genetic heterogeneity.

Results

Participation rates

Overall, the NTR has enjoyed excellent participation rates. The response rate for age 3 was 75.5% of all families (registered with the NTR) that were contacted when the twins turned 3 years old. This number excludes both 'true' non-responders and families who changed domicile without giving notice. We calculated the continued participation rate by considering complete overlapping birth cohorts for adjoining ages. For example, data from the 1986, 1987, and 1988 birth cohorts were used to calculate the number of maternal reports collected when the twins were age 10 and age 12. Continued participation rate was about 80%, for all age-intervals and in each zygosity group. Despite the large longitudinal response rate, we cannot exclude the possibility that the loss of participants at follow-up assessments was selective. Testing revealed that non-participation at age 7, age 10, and age 12 was positively related to the twins' overactive problem behavior at age 3 by maternal report. At age 7 and age 12, the effect was small but significant (age 7, $F = 11.79$, $p < .01$; age 12, $F = 5.83$, $p = .02$). At age 7, the mean OA scores at age 3 of participants and non-participants were 2.67 and 2.86, respectively. At age 12, the mean OA scores at age 3 of participants and non-participants were 2.59 and 2.81, respectively. This result suggests that the missing data are not missing completely at random. However, the hypothesis that the data are missing at random remains tenable. About half of the mothers, who did not return the questionnaire when the twins were 7, 10, or 12 years of age (i.e., the sample used for the calculation of the

selection bias), participated at least once at another measurement occasion. If a high OA score at age 3 had a direct influence on the mothers' decision to refuse participation at a follow-up age, one would expect the mother to leave the study entirely beyond the twins' age of 3 years. Given that this was not the case in 50% of the mothers suggests that the OA score at age 3 plays a modest role in explaining why mothers dropped out of the study. Although not conclusive, these results support the hypothesis that the data are missing at random.

Ratings of behavior and phenotypic stability

Table 2 contains the OA and AP scores by gender and age.

Boys had higher mean OA and AP scores than girls ($\chi^2_4 = 427.59$, $p < .0001$) at all ages. From a developmental perspective, the changes over age are most interesting. Since OA at age 3 and AP at older ages differ with respect to item content, we did not consider the change in mean scores in this interval. The different pattern from age 3 to age 7 between boys and girls is striking, however. The age interval from 7 to 10 years appears to be a period of minimal change in mean AP, indicated by a non-significant effect for girls and a small significant effect for boys ($\chi^2_1 = 4.61$, $p = .03$). In both girls and boys, the decrease in mean AP from age 10 to age 12 was significant (boys, $\chi^2_1 = 36.43$, $p < .001$; girls, $\chi^2_1 = 37.34$, $p < .001$).

Table 3 contains the inter-occasion correlations.

Two interesting results emerge from Table 3. First, both boys and girls display a moderate degree of stability from ages 3 to 7, and a much larger degree of stability from ages 7 to 10, and from ages 10 to 12. Second, the pattern in correlations is consistent across gender. The correlations between boys and girls do not differ significantly ($\chi^2_6 = 6.66$, $p > .05$). This implies that although boys have higher OA and AP scores, the observed stability across age is equal for boys and girls.

From a clinical perspective, it is of interest to focus on children with a high score on the AP scale. Several studies reported close convergence between high scores on the AP scale and ADHD by DSM criteria (Biederman et al., 1993; Chen, Faraone, Biederman, & Tsuang, 1994; Hudziak et al., in review; Kasius, Ferdinand, Van den Berg, & Verhulst, 1997; Steingard, Biederman, Doyle, & Sprich-Buckminster, 1992). To compare between CBCL AP and DSM ADHD, prevalence rates were calculated by imposing a cut-off score on the distribution of scores of the twins. The cut-off scores were derived from a Dutch community sample (age 3; Koot, 1993) and a norm sample (ages 7, 10, 12; Verhulst, Van der Ende, & Koot, 1996). This cut-off score corresponds to a *T*-score of 67. Children with a larger *T*-score are considered to be borderline or clinical cases (Achenbach, 1991). In taking the Dutch community and norm samples as a reference, we assume that the twin distribution resembles the distribution of scores of the singletons in those samples. The resulting prevalence rates are 3.1%, 6.3%, 7.0%, 5.2% for boys and 2.1%, 5.2%, 5.0%, 3.6% for girls, at ages 3, 7, 10, and 12 years, respectively. Prevalence rates for deviant AP are comparable to the norm rates (Verhulst et al., 1996), but prevalence rates for deviant OA reported here are modest compared to the community rates (Koot, 1993). Slightly more than 400 children participated in the community sample. This relatively small sample, and the inclusion of 10 clinical cases in the community sample, possibly explains part of the discrepancy in prevalence rates between the twin and community sample. In a follow-up study of the community sample (Mesman & Koot, 2001) it was found that the prevalence of DSM-IV diagnoses at ages 10 and 11 was somewhat higher compared to other community reports. Maternal age may further explain the lower number of borderline and clinical cases for OA in the twin sample. Koot (1993) found that age of the mother was a protective factor in the degree of problem behavior displayed by children (see also, Orlebeke, Knol, Boomsma, & Verhulst, 1998). Given that DZ twinning rate increases with age of the mother, the entire twin sample may have been rated by mothers who are on average older than the mothers who participated in the community study. In the present sample DZ twins have older mothers than MZ twins ($t = 7.59$, $df = 5379$, $p < .0001$), and male DZ twins have fewer OA problems than male MZ twins ($\chi^2_1 = 7.46$, $p < .01$). Linear regression revealed a small but significant contribution of maternal age to OA ($R^2 = .025$, $p < .05$). These results suggest that with increasing age, mothers report less problem behavior in their offspring. However, it remains uncertain

Table 2 Means and standard deviation (sd) for OA at age 3 and AP at ages 7, 10 and 12

	Boys			Girls		
	N	Mean	Sd	*N*	Mean	Sd
Age 3	5346	2.92	2.22	5520	2.48 *	2.15
Age 7	5000	3.37	2.95	5211	2.40 *	2.59
Age 10	2866	3.47 #	3.18	3037	2.38 *	2.58
Age 12	1433	3.09 #	2.94	1486	2.06 * #	2.41

Note: Significance was established at $p < .05$. N = number of individual twins. * = mean score of girls differs significantly from mean score of boys. # = mean score differs significantly from mean score at previous age within gender.

Table 3 Maximum likelihood within-person phenotypic correlations by gender. Correlations for boys and girls are reported below and above diagonal, respectively

Boys\Girls	Age 3	Age 7	Age 10	Age 12
Age 3	1.00	.41	.37	.38
Age 7	.41	1.00	.68	.65
Age 10	.38	.69	1.00	.72
Age 12	.35	.67	.75	1.00

whether maternal age explains the difference in prevalence rates *between* studies.

To establish the stability of OA and AP scores within the normal range, we calculated the percentage of boys and girls for whom a diagnosis of DSM ADHD most likely never applies (*T*-score < 67). Around 95% of the children who fall within the normal range of OA at age 3 do not experience AP at age 7. For the following age intervals, from 7 to 10 years and from 10 to 12 years, the percentage of children whose AP score continued to fall below the borderline cut-off score remained at 95%. No effect of gender was observed.

Heritability of OA and AP

Behavior genetic analyses of the data started with the calculation of twin correlations. With longitudinal data, correlations are calculated both within each age and across two ages (cross-correlations). The first give insight in the role of genes and environment at a specific age. The latter are informative with respect to the importance of genes and environment in explaining stability over time. Twin correlations and cross-correlations did not differ between boys and girls and between same-sex dizygotic pairs and opposite-sex dizygotic pairs. Therefore, the within and across age correlations were calculated for all MZ and all DZ twins.

Inspection of the difference between MZ and DZ twin within age correlations (on the diagonal in Table 4) suggests that variance in OA and AP is explained by additive genetic, genetic dominance, and unique environmental effects. MZ cross-correlations are also larger than DZ cross-correlations (see off-diagonals in Table 4). Inspection of the variances reveals that at age 3, DZ variances exceed MZ variances both in girls and boys. The discrepancy in variances is consistent with the presence of a contrast effect.

Table 5 contains an overview of the results of the longitudinal analyses. Model 1 serves as the baseline model. The importance of genetic dominance (Model 2) and the contrast effects (Model 3) at ages 7, 10 and 12 were evaluated against this model. The chi-square test based on difference in log-likelihood and associated degrees of freedom between Model 1 and Model 2 suggested that genetic dominance might be omitted from the model. However, we know from power studies (Eaves, 1972; Posthuma & Boomsma, 2000; Rietveld et al., 2003b) that the current twin sample does not provide much power to detect genetic dominance and it was therefore decided to retain genetic dominance in the model at this stage of model fitting. Estimates provided by Model 1 indicated a larger contrast effect at age 3 (–.08) compared to the contrast effect at older ages (ranging from .02 to .05). The non-significant difference between Model 1 and Model 3 suggested that the three contrast parameters as specified at ages 7, 10, and 12 are negligible. The chi-square based on the difference in log-likelihood of Models 3 and 4 suggests that the contrast parameter at age 3 is significant. We next fitted Model 5 to assess the importance of genetic dominance in explaining observed variation for AP at ages 7, 10, and 12. As indicated by the significant chi-square test, genetic dominance appeared a significant source of variance at these ages. We therefore accepted Model 3 as the best fitting model.

The decomposition of phenotypic variance based on Model 3 is given in the top part of Table 6. The diagonals contain the percentages for boys and girls, respectively. The off-diagonals are the relative contributions of A, D, and E to the observed covariance over time. Estimates for boys and girls are listed in the lower and upper triangular, respectively. These percentages indicate the extent to which the observed covariance of OA with AP from age to age is due to additive genetic, genetic dominance, and unique environmental influences. These results are corrected for the contrast effects. Generally, these results are similar to the results of other twin studies. Genetic and environmental correlations are listed in the lower part of Table 6.

As listed on the diagonals, around one-third to half of the observed variation in OA and AP at each age is explained by additive genetic influences. In

Table 4 Within age twin correlations (diagonal: MZ and DZ correlations) and across age cross-twin correlations (e.g., OA-3-years in twin 1 with AP-7-years in twin 2). MZ correlations below and DZ correlations above diagonal

MZ\DZ	Age 3	Age 7	Age 10	Age 12
Age 3	.66\.13	.13	.13	.12
Age 7	.34	.71\.28	.19	.18
Age 10	.33	.54	.72\.28	.19
Age 12	.32	.52	.55	.72\.26

Table 5 Longitudinal model fitting results by Cholesky decomposition

	Model	Age	–2 LL	*df*	Comparison	χ^2	*df*	*p*-value
1.	ADE + b	all ages	68064.85	29791				
2.	AE + b	all ages	68092.00	29811	1.	27.15	20	.13
3.	ADE + b	age 3	68069.45	29797	1.	4.60	6	.40
	ADE	age 7, 10, 12						
4.	ADE	all ages	68082.38	29799	3.	12.93	2	.00
5.	ADE + b	age 3	68107.31	29815	4.	37.86	18	.00
	AE	age 7, 10, 12						

Table 6 *Top part* includes percentages of total variances (diagonal) and covariances (off-diagonal) explained by additive genetic, genetic dominance, and unique environmental components based on best fitting models. Percentages for boys and girls are reported below and above diagonal, respectively. *Lower part* includes correlations calculated for additive genetic, genetic dominance, and unique environmental sources of variance between different ages. Correlations for boys and girls are reported below and above diagonal, respectively

Relative proportions of variance and covariance

	A %				D %				E %			
Boys\Girls	OA 3	AP 7	AP 10	AP 12	OA 3	AP 7	AP 10	AP 12	OA 3	AP 7	AP 10	AP 12
OA 3	50\41	73	79	75	22\33	17	13	14	28\26	10	8	11
AP 7	59	33\57	50	53	31	39\16	31	28	10	28\27	19	19
AP 10	86	31	41\48	47	6	51	31\25	32	8	18	28\27	21
AP 12	71	24	31	40\54	16	55	45	30\18	13	21	24	30\28

Correlations between different ages

	A				D				E			
Boys\Girls	OA 3	AP 7	AP 10	AP 12	OA 3	AP 7	AP 10	AP 12	OA 3	AP 7	AP 10	AP 12
OA 3	1.00	.60	.66	.57	1.00	.30	.16	.20	1.00	.15	.12	.14
AP 7	.57	1.00	.62	.57	.41	1.00	.99	1.00	.15	1.00	.46	.41
AP 10	.68	.56	1.00	.61	.08	.94	1.00	1.00	.11	.42	1.00	.50
AP 12	.49	.42	.53	1.00	.20	.98	.99	1.00	.14	.45	.58	1.00

girls, additive genetic influences are more important in explaining variation in AP than in boys. With the exception of AP measured in boys at age 7 (A 33%, D 39%), genetic dominance is found to contribute less to the observed variation at each age. Averaged over age and gender, broad heritability (the percentage of variance due to additive and dominance effects) is between 70% and 74%. The residual variance is explained by the unique (unshared) environment.

In both boys and girls, the covariance between OA and AP is explained by additive genetic influences to a much larger extent compared to the covariance among the three measurements of AP. As in the age-specific results, additive genetic effects are important in girls whereas non-additive effects are important in boys. Environmental influences that are unshared by family members contribute little to stability over time (ranging from 8% to 24%).

The correlations in the lower part of Table 6 indicate the degree of overlap between influences at one age and influences at subsequent ages. The additive genetic correlations are estimated between .42 and .68, and thus suggest only partial overlap in additive genetic effects. Correlations genetic dominance effects range from .94 to 1.00 over ages 7, 10, and 12. These correlations suggest a high degree of stability of dominance effects. However, the genetic dominance correlations between OA at age 3 and AP at later ages are much smaller (.08 to .41). A similar pattern of correlations is observed in the case of the unique environmental effects. The environmental correlations between OA and AP are estimated around .13. This is much lower than the environmental correlations (about .46) between AP measured at ages 7, 10, and 12 years.

Discussion

We analyzed maternal ratings of Overactivity (OA) and Attention Problems (AP) in twins at ages 3, 7, 10, and 12 years. The longitudinal allowed the study of the development of problem behavior over this period. We addressed the following questions. To what degree do genetic and environmental influences contribute to variation in OA and AP? Do results vary with the age or gender of the child? How do genetic and environmental influences contribute to the observed stability in OA and AP? Do results depend on the age-interval being studied? Is observed stability in boys and girls explained by an identical pattern of genetic and environmental influences? To what degree do boys and girls share these influences? Age-specific and longitudinal results are discussed at the phenotypic level first, followed by a discussion of genetic and environmental results.

Gender differences in mean scores and prevalences for OA and AP

Studies on OA and AP and other ADHD-related behaviors in girls are lacking (Gaub & Carlson, 1997; Klein & Mannuzza, 1991). In this paper we present data of large samples of twin families, with an equal number of girls and boys. The well-known gender difference with boys displaying more OA and AP was observed at each age. Even at the age of 3, boys display more OA problems than girls. Clinical studies have indicated that severe problem behavior can be identified in very young children (see for review, Campbell, 1995; Keenan & Wakschlag, 2000; Shaw, Owens, Giovannelli, & Winslow, 2001) and that the onset of ADHD is during the pre-school period (Barkley, Fisher, Edelbrock, & Smallish, 1990;

Kadesjö, Kadesjö, Hägglöf, & Gillberg, 2001; McGee, Williams, & Feehan, 1992). Children who suffer from attention deficits, impulsivity, or hyperactivity often display more problematic behavior when they enter the school system. This increase may reflect a true rise in problem behavior, a change in rating behavior of the mother, or a combination of both. Classroom discipline requires children to sit still and pay attention to the teacher. The inability of children suffering from OA and AP to do so may elicit negative reactions from teachers and peers, which in turn may affect the degree of problem behavior. The interaction between the school-age child and others may also lead to an increased awareness in the mother of the degree of AP that her child displays. In our study we observe high mean AP scores in boys at the age of 7 and 10 years. A decline of AP in boys is observed at age 12. Girls display a small but steady decline in maternally reported problem behavior from an earlier age. The difference between boys and girls is also seen at the higher end of the distribution. At every age, the ratio approaches 1.5 to 1, which implies that only a small proportion of boys outnumber girls on DSM criteria in the present study. This gender ratio is not exceptional in non-clinical samples (Biederman, 1998; Pineda et al., 1999; Szatmari, Offord, & Boyle, 1989). In a recent population study of hyperactivity in 1200 Dutch children aged 2 to 11 years (Centraal Bureau voor de Statistiek, 2002b), a similar gender ratio was observed.

Developmental changes from OA to AP

The observed age-specific gender differences in means notwithstanding, the degree of stability in boys and girls is about equal. We observed moderate stability of OA at age 3 to AP at age 7 (correlation .41), and greater stability of AP from age 7 to age 10, and from age 10 to age 12 (correlations about .70). Possible explanations for the moderate degree of stability include the amount of time between assessments, developmental changes, and the change in phenotype (OA at age 3, AP thereafter). The interval between ages 3 and 7 is not smaller than the interval between ages 7 and 12. Given that the observed correlation between ages 7 and 12 is comparable to those between ages 7 and 10, and ages 10 and 12, time interval does not seem to be important in explaining the difference in stability. Developmental change may be a factor, as findings from neurological and neuropsychological studies suggest that the age period from the preschool and early school years is marked by the acquisition of attention skills (Bennett Murphy, Murphy, & Rose, 2001; Woody-Ramsey & Miller, 1988). Thus, 3-year-old children may suffer from OA problems relative to their developmental maturity, but by age 7 years, the majority of children should have progressed through this developmental period and thus have relatively stable attention skills. Any deficits may be more likely to persist. Another reason for the moderate stability may be the differences in assessment instrument. At age 3, the OA scale of the CBCL/2-3 is composed of 5 items. At older ages, the AP scale of the CBCL/4-18 is composed of 11 items. As can be seen in Appendix I, OA and AP share two items (item 5 (OA) with item 8 (AP); item 6 (OA) with item 10 (AP)). If the moderate degree of stability from OA to AP is due to developmental change, the correlation of the overlapping items between ages 3 and 7 should resemble those for the total scores on OA to AP. Specifically, because age is hypothesized to be very relevant and item content is identical, item stability from ages 3 to 7 should approach the observed moderate OA–AP stability from age 3 and beyond (around .40) and should not approach the observed large AP–AP stability from age 7 and beyond (around .70). The observed correlations ranged from .27 to .44, i.e., they resemble the correlation between OA and AP. We therefore suggest that the modest association between OA at age 3 and AP at age 7 represents developmental change. Hart, Lahey, Loeber, Applegate, and Frick (1995) also reported that the decline in hyperactive behavior in a clinical sample was related to the increasing age of the subjects. Artifacts of repeated assessment by the same informant and passage of time between assessments were found to be less plausible in explaining the decrease in hyperactivity.

Our results are comparable with those reported in other studies that used the CBCL. Achenbach and Rescorla (2000) reported a correlation of .40 between AP at age 3 (CBCL/1.5-5; revision of the CBCL/2-3) and AP at age 7. Three out of five items of the Dutch OA scale are included in the AP scale derived in analyses of the CBCL/1.5-5. Koot (1993) reported a correlation of .47 between OA at ages 2–3 years and AP at ages 4–5 years. The high degree of AP stability found here is in line with those reported for repeatedly assessed children varying in age from 7 to 10 years (Verhulst et al., 1995). In a sample of adolescent adoptees, Van der Valk et al. (1998) observed a correlation of around .65 between two assessments that covered a three-year interval.

Age-specific heritability estimates

Individual differences in childhood OA and AP are mainly due to individual differences in genetic factors. The contribution of unique (or unshared) environmental effects to individual differences in OA and AP is considerably smaller. The genetic component also includes any effects of an interaction between shared family environment and genotype (Boomsma & Martin, 2002; Plomin, DeFries, & Loehlin, 1977). Imagine two children with a different score on the AP scale. The children may behave in a similar way when the environment contains only minimal sources of distraction. When the children enter a situation with

many stimuli, they may differ in their response to this new situation. In the classical twin design, it is not possible to distinguish between variance due to the interaction of shared environment and genotype and variance due to genetic effects only. An interaction between genetic effects and unique environmental effects contributes to the unique environmental variance estimate. Because the unique environment was estimated at around a modest 25%, this type of interaction, if present, is unlikely to be large.

Overall, our results are in line with the results of other twin studies of AP (Edelbrock et al., 1995; Gjone et al., 1996; Hudziak et al., 2000; Schmitz et al., 1996; Zahn-Waxler et al., 1996) and phenotypes related to ADHD (Eaves et al., 1997; Kuntsi et al., 2000; Martin et al., 2002; Nadder et al., 1998; Price et al., submitted; Rhee et al., 1999; Sherman et al., 1997; Thapar et al., 1995; Thapar et al., 2000). In many of these studies gender and age effects were not addressed. Here, we found that heritability estimates are equal for boys and girls, although boys display significantly more variation in their behavior at ages 10 and 12. Further, genetic heterogeneity was found to be absent, which suggests that the same set of genes contribute to individual differences in boys and girls.

At age 3, a significant contrast effect was detected. Given the results of other twin studies that included multiple raters (e.g., Simonoff et al., 1998), it is likely that the contrast effect is due to a bias in the ratings of the mother. This bias is due to the mothers comparing the twins with one another and overestimating any perceived differences. As children get older and mothers are exposed to the behavior of children other than their own twins, they rely less on the comparison of the twins in judging the children's behavior. In other words, during children's development, the mothers' frame of reference may change from children inside the home to children outside the home. This would account for the absence of a contrast effect at ages 7, 10, and 12.

Genetic and environmental stability of OA and AP

Additive genetic effects were found to make the greatest contribution to the stability between OA age 3 and AP beyond age 3. Genetic dominance and unique environmental effects contributed relatively little. The differences in these contributions to stability continued over the intervals including ages 7, 10, and 12. After age 3, additive genetic effects were found to be more important in girls and genetic dominance more important in boys. The genetic correlations between the ages suggest that additive genetic effects are far from perfectly stable. Only a subset of genes that operates at one age does so at a later age. This result is consistent over the age-intervals and over gender. In contrast, genetic dominance and the unique environment effects varied with age-interval. Very low correlations are observed for the intervals that include age 3. This indicates that the observed correlation between OA (age 3) and AP (beyond age 3) is nearly entirely due to additive genes. A more complex picture arises for the intervals including ages 7, 10, and 12. When all genetic and environmental correlations are taken into account, it emerges that the observed stability of AP is stability of additive genetic, dominance, and environmental effects common to these 3 ages. These results agree with results reported by Schmitz et al. (submitted).

Limitations of the study

We found that the unique environmental effects explained a quarter of the observed variance in OA and AP at each age. As we did not include an actual measure of the environment, we cannot identify any specific aspect of unique environment that contributes to individual differences in OA and AP, except to say that this 25% also includes measurement error. The nature of these effects is thus partly unknown.

The interval between ages 3 and 7 is large, and possible developmental changes may not have been captured by this study. This seems unlikely for boys, given the sharp increase in problem behavior during this age period. However, girls may first have increased in problem behavior shortly after age 3 and then decreased by the time they were 7 years old.

Neither OA nor AP is a measure of ADHD. Here we used AP as a marker for attention problem behavior, which is an important aspect of ADHD. AP has been shown to be predictive of DSM-III-R ADHD (Chen et al., 1994) and DSM-IV ADHD (Hudziak et al., in review). At present, we are collecting DSM-IV data. These data will enable us to explore the exact relation between OA, AP, and ADHD in the present sample.

Conclusions

Our study supports the findings reported earlier in smaller, cross-sectional studies, that genetic influences are important in the development of AP. The longitudinal data revealed that genetic influences are also important for the stability of these problems. Furthermore, we found that children who do not have AP at one stage of development are unlikely to develop AP at a later stage.

Acknowledgements

The work of the first author is supported by grant 96/22 from the Universitair Stimulerings Fonds to D.I. Boomsma. The work by the second author is supported by grant MH52813 from NIMH to J.J. Hudziak. We thank Professor Hans Koot for kindly providing his CBCL dataset on 3-year-old singletons (referred to in text as the community sample).

Correspondence to

M.J.H. Rietveld, Vrije Universiteit, Department of Biological Psychology, Van der Boechhorststraat 1, 1081 BT Amsterdam, The Netherlands; Email: mjh.rietveld@psy.vu.nl

References

Achenbach, T.M. (1991). *Manual for the Child Behavior Checklist/4-18*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1992). *Manual for the Child Behavior Checklist/2-3*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M., & Rescorla, L.A. (2000). *Manual for the ASEBA Preschool Forms and Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth and Families.

Barkley, R.A., Fisher, M., Edelbrock, C.S., & Smallish, L. (1990). The adolescent outcome of hyperactive children diagnosed by research criteria: I. An 8-year prospective follow-up study. *Journal of the American Academy of Child and Adolescence Psychiatry*, *29*, 546–557.

Bennet Murphy, L., Murphy, C.E., & Rose, C.L. (2001). Sustained attention and unintentional injury among preschool-aged children. *Child Neuropsychology*, *7*, 72–83.

Biederman, J. (1998). Attention-Deficit/Hyperactivity Disorder: A life-span perspective. *Journal of Clinical Psychiatry*, *59*, 4–16.

Biederman, J., Faraone, S.V., Doyle, A., Lehman, B.K., Kraus, I., Perrin, J., & Tsuang, M.T. (1993). Convergence of the Child Behavior Checklist with structured interview-based psychiatric diagnoses of ADHD children with and without comorbidity. *Journal of Child Psychology and Psychiatry*, *34*, 1241–1251.

Biederman, J., Mick, E., & Faraone, S.V. (2000). Age-dependent decline of symptoms of Attention Deficit Hyperactivity Disorder: Impact of remission definition and symptom type. *American Journal of Psychiatry*, *157*, 816–818.

Boomsma, D.I. (1998). Twin registers in Europe: An overview. *Twin Research*, *1*, 34–51.

Boomsma, D.I., & Martin, N.G. (2002). Gene–environment interactions. In H. D'Haenen, J. A. den Boer, & P. Willner (Eds.), *Biological psychiatry* (pp. 181–187). Chichester: John Wiley & Sons, Ltd.

Boomsma, D.I., & Molenaar, P.C.M. (1986). Using LISREL to analyze genetic and environmental covariance structure. *Behavior Genetics*, *16*, 237–250.

Campbell, S.B. (1995). Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology and Psychiatry*, *36*, 113–149.

Carey, G. (1986). Sibling imitation and contrast effects. *Behavior Genetics*, *16*, 319–341.

Centraal Bureau voor de Statistiek. (2002a). *Enquete Beroepsbevolking 1999. Standaard Beroepsclassificatie 1992*. Heerlen: Centraal Bureau voor de Statistiek. [information obtained from the internet address http://www.cbs.nl/].

Centraal Bureau voor de Statistiek. (2002b). *Permanent Onderzoek Leefsituatie, module Gezondheid en Arbeid, specifieke gezondheidsmetingen bij kinderen 2001*. Heerlen: Centraal Bureau voor de Statistiek. [information obtained from the internet address http://www.cbs.nl/].

Chen, W.J., Faraone, S.V., Biederman, J., & Tsuang, M.T. (1994). Diagnostic accuracy of the Child Behavior Checklist scales for attention-deficit hyperactivity disorder: A receiver-operating characteristic analysis. *Journal of Consulting and Clinical Psychology*, *62*, 1017–1025.

De Groot, A., Koot, H.M., & Verhulst, F.C. (1994). Cross-cultural generalizability of the Child Behavior Checklist cross-informant syndromes. *Psychological Assessesment*, *6*, 225–230.

Eaves, L.J. (1972). Computer simulation of sample size and experimental design in human psychogenetics. *Psychological Bulletin*, *77*, 144–152.

Eaves, L.J. (1976). A model for sibling effects in man. *Heredity*, *36*, 205–214.

Eaves, L.J., Rutter, M., Silberg, J.L., Shillady, L., Maes, H.H., & Pickles, A. (2000). Genetic and environmental causes of covariation in interview assessments of disruptive behavior in child and adolescent twins. *Behavior Genetics*, *30*, 321–334.

Eaves, L.J., Silberg, J.L., Maes, H.H., Simonoff, E., Pickles, A., Rutter, M., Neale, M.C., Reynolds, C.A., Erikson, M.T., Heath, A.C., Loeber, R., Truett, K.R., & Hewitt, J.K. (1997). Genetics and developmental psychopathology: 2. The main effects of genes and environment on behavioral problems in the Virginia Twin Study of Adolescent Behavioral Development. *Journal of Child Psychology and Psychiatry*, *38*, 965–980.

Edelbrock, C., Rende, R., Plomin, R., & Thompson, L.A. (1995). A twin study of competence and problem behavior in childhood and early adolescence. *Journal of Child Psychology and Psychiatry*, *36*, 775–785.

Falconer, D.S., & Mackay, T.F.C. (1996). *Introduction to quantitative genetics* (4th edn.). Harlow: Longman Group Ltd.

Gaub, M., & Carlson, C.L. (1997). Gender differences in ADHD:A meta-analysis and critical review. *Journal of American Academy of Child and Adolescent Psychiatry*, *36*, 1036–1045.

Gjone, H., Stevenson, J., & Sundet, J.M. (1996). Genetic influence on parent-reported attention-related problems in a Norwegian general population twin sample. *Journal of American Academy of Child and Adolescent Psychiatry*, *35*, 588–596.

Hart, E.L., Lahey, B.B., Loeber, R., Applegate, B., & Frick, P.J. (1995). Developmental change in Attention-Deficit Hyperactivity Disorder in boys: A four-year longitudinal study. *Journal of Abnormal Child Psychology*, *23*, 729–747.

Hudziak, J.J. (2001). The role of phenotypes (diagnoses) in genetic studies of attention-deficit/hyperactivity disorder and related child psychopathology. *Child and Adolescent Psychiatric Clinics of North America*, *10*, 279–297.

Hudziak, J.J., Heath, A.C., Madden, P.A.F., Reich, W., Bucholz, K.K., Slutske, W.S., Bierut, L.J., Neuman, R.J., & Todd, R.D. (in review). The genetic analysis of parental reports of DSM-IV. *Journal of American Academy of Child and Adolescent Psychiatry*.

Hudziak, J.J., Rudiger, L.P., Neale, M.C., Heath, A.C., & Todd, R.D. (2000). A twin study of inattentive, aggressive and anxious/depressed behaviors. *Journal of American Academy of Child and Adolescent Psychiatry*, *39*, 469–476.

Kadesjö, C., Kadesjö, B., Hägglöf, B., & Gillberg, C. (2001). ADHD in Swedish 3- to 7-year-old children. *Journal of American Academy of Child and Adolescent Psychiatry*, *40*, 1021–1028.

Kasius, M.C., Ferdinand, R.F., Van den Berg, H., & Verhulst, F.C. (1997). Associations between different diagnostic approaches for child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, *38*, 625–632.

Keenan, K., & Wakschlag, L.S. (2000). More than the terrible twos: The nature and severity of behavior problems in clinic-referred preschool children. *Journal of Abnormal Child Psychology*, *28*, 33–46.

Klein, R.G., & Mannuzza, S. (1991). Long-term outcome of hyperactive children: A review. *Journal of American Academy of Child and Adolescent Psychiatry*, *30*, 383–387.

Koot, H.M. (1993). *Problem Behavior in Dutch Preschoolers*. Doctoral dissertation, Erasmus University, Rotterdam.

Koot, H.M., Van den Oord, E.J.C.G., Verhulst, F.C., & Boomsma, D.I. (1997). Behavioural and emotional problems in young pre-schoolers: Cross-cultural testing of the validity of the Child Behavior Checklist 2/3. *Journal of Abnormal Child Psychology*, *25*, 183–196.

Kuntsi, J., Gayán, J., & Stevenson, J. (2000). Parents' and teachers' ratings of problem behaviours in children: Genetic and contrast effects. *Twin Research*, *3*, 251–258.

Little, R.J.A., & Rubin, D.B. (1987). *Statistical analysis with missing data*. New York: Wiley.

Loehlin, J.C. (1992). *Latent variable models: An introduction to factor, path, and structural analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Martin, N.G., & Eaves, L.J. (1977). The genetical analysis of covariance structure. *Heredity*, *38*, 79–95.

Martin, N., Scourfield, J., & McGuffin, P. (2002). Observer effects and heritability of childhood attention-deficit hyperactivity disorder symptoms. *British Journal of Psychiatry*, *180*, 260–265.

McGee, R., Williams, S., & Feehan, M. (1992). Attention Deficit Disorder and age of onset of problem behaviors. *Journal of Abnormal Child Psychology*, *20*, 487–502.

Mesman, J., & Koot, H.M. (2001). Early preschool predictors of preadolescent internalizing and externalizing DSM-IV diagnoses. *Journal of American Academy of Child and Adolescent Psychiatry*, *40*, 1029–1036.

Nadder, T.S., Silberg, J.L., Eaves, L.J., Maes, H.H., & Meyer, J.M. (1998). Genetic effects on ADHD symptomatology in 7- to 13-year-old twins: Results from a telephone survey. *Behavior Genetics*, *28*, 83–100.

Nadder, T.S., Silberg, J.L., Rutter, M., Maes, H.H., & Eaves, L.J. (2001). Comparison of multiple measures of ADHD symptomatology: A multivariate genetic analysis. *Journal Child Psychology and Psychiatry*, *42*, 475–486.

Neale, M.C. (1997). *Mx: Statistical modeling*. Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Neale, M.C., Cardon, L.R., & The North Atlantic Treaty Organization. Scientific Affairs Division. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic Publishers.

Neale, M.C., & Stevenson, J. (1989). Rater bias in the EASI Temperament Scales: A twin study. *Journal of Personality and Social Psychology*, *56*, 446–455.

Orlebeke, J.F., Knol, D.L., Boomsma, D.I., & Verhulst, F.C. (1998). Frequency of parental report of problem behavior in children decreases with increasing maternal age of delivery. *Psychological Reports*, *82*, 395–404.

Pineda, D., Ardila, A., Rosselli, M., Arias, B.E., Henao, G.C., Gomez, L.F., Mejia, S.E., & Miranda, M.L. (1999). Prevalence of Attention-Deficit/Hyperactivity Disorder symptoms in 4- to 17-year-old children in the general population. *Journal of Abnormal Child Psychology*, *27*, 455–462.

Plomin, R., DeFries, J.C., & Loehlin, J.C. (1977). Genotype–environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*, 309–322.

Plomin, R., DeFries, J.C., McClearn, G.E., & McGuffin, P. (2001). *Behavior genetics* (4th edn). New York: W.H. Freeman.

Posthuma, D., & Boomsma, D.I. (2000). A note on the statistical power in extended twin designs. *Behavior Genetics*, *30*, 147–158.

Price, T.S., Simonoff, E., Waldman, I., Asherson, P., Curran, S., & Plomin, R. (submitted). What is stable about hyperactive behaviors in pre-school children is genetic: Implications for molecular genetic studies.

Rhee, S.H., Waldman, I.D., Hay, D.A., & Levy, F. (1999). Sex differences in genetic and environmental influences on DSM-III-R Attention-Deficit/Hyperactivity Disorder. *Journal of Abnormal Psychology*, *108*, 24–41.

Rietveld, M.J.H., Hudziak, J.J., Bartels, M., Van Beijsterveldt, C.E.M., & Boomsma, D.I. (2003a). Heritability of attention problems in children: I. Cross-sectional results from a study of twins, age 3 to age 12. *American Journal of Medical Genetics B*, *117b*, 102–113.

Rietveld, M.J.H., Posthuma, D., Dolan, C.V., & Boomsma, D.I. (2003b). ADHD: Sibling interaction or dominance: An evaluation of statistical power. *Behavior Genetics*, *33*, 247–255.

Rietveld, M.J.H., Van der Valk, J.C., Bongers, I.L., Stroet, T.M., Slagboom, P.E., & Boomsma, D.I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Research*, *3*, 134–141.

Ross, D.M., & Ross, S.A. (1982). *Hyperactivity. Current issues, research, and theory*. New York: John Wiley & Sons, Inc.

Rutter, M., Giller, H., & Hagell, A. (1998). *Antisocial behavior by young people*. New York/Cambridge: Cambridge University Press.

Saudino, K.J., Cherny, S.S., & Plomin, R. (2000). Parent ratings of temperament in twins: Explaining the 'too low' DZ correlations. *Twin Research*, *3*, 224–233.

Schmitz, S., Corley, R.P., Hewitt, J.K., Zahn-Waxler, C., Emde, R.N., & DeFries, J.C. (submitted). Attention problems and emotionality at ages 7 and 12.

Schmitz, S., Fulker, D.W., & Mrazek, D.A. (1995). Problem behavior in early and middle childhood: An initial behavior genetic analysis. *Journal of Child Psychology and Psychiatry*, *36*, 1443–1458.

Schmitz, S., Saudino, K.J., Plomin, R., Fulker, D.W., & DeFries, J.C. (1996). Genetic and environmental influences on temperament in middle childhood: Analyses of teacher and tester ratings. *Child Development*, *67*, 409–422.

Shaw, D.S., Owens, E.B., Giovannelli, J., & Winslow, W.B. (2001). Infant and toddler pathways leading to early externalizing disorders. *Journal of American Academy of Child and Adolescent Psychiatry*, *40*, 36–43.

Sherman, D.K., Iacono, W.G., & McGue, M.K. (1997). Attention-deficit hyperactivity disorder dimensions: A twin study of inattention and impulsivity-hyperactivity. *Journal of American Academy of Child and Adolescent Psychiatry*, *36*, 745–753.

Simonoff, E., Pickles, A., Hervas, A., Silberg, J.L., Rutter, M., & Eaves, L. (1998). Genetic influences on childhood hyperactivity: Contrast effects imply parental rating bias, not sibling interaction. *Psychological Medicine*, *28*, 825–837.

Steingard, R., Biederman, J., Doyle, A., & Sprich-Buckminster, S. (1992). Psychiatric comorbidity in attention deficit disorder: Impact on the interpretation of Child Behavior Checklist results. *Journal of American Academy of Child and Adolescent Psychiatry*, *31*, 449–454.

Szatmari, P., Offor, D.R., & Boyle, M.H. (1989). Ontario Child Health Study: Prevalence of Attention Deficit Disorder with Hyperactivity. *Journal of Child Psychology and Psychiatry*, *30*, 219–230.

Thapar, A., Harrington, R., Ross, K., & McGuffin, P. (2000). Does the definition of ADHD affect heritability? *Journal of American Academy of Child and Adolescent Psychiatry*, *39*, 1528–1536.

Thapar, A., Hervas, A., & McGuffin, P. (1995). Childhood hyperactivity scores are highly heritable and show sibling competition effects: Twin study evidence. *Behavior Genetics*, *25*, 537–544.

Van den Oord, E.J.C.G. (1993). *A genetic study of problem behaviors in children*. Doctoral dissertation, Erasmus University, Rotterdam.

Van den Oord, E.J.C.G., Verhulst, F.C., & Boomsma, D.I. (1996). A genetic study of maternal and paternal ratings of problem behaviors in 3-year-old twins. *Journal of Abnormal Psychology*, *105*, 349–357.

Van der Valk, J.C., Verhulst, F.C., Neale, M.C., & Boomsma, D.I. (1998). Longitudinal genetic analysis of problem behaviors in biologically related and unrelated adoptees. *Behavior Genetics*, *28*, 365–380.

Verhulst, F.C., Van der Ende, J., & Koot, H.M. (1996). *Handleiding voor de CBCL/4-18*. Rotterdam: Afdeling Kinder- en Jeugdpsychiatry, Sophia Kinderziekenhuis / Academisch Ziekenhuis Rotterdam / Erasmus Universiteit Rotterdam.

Verhulst, F.C., & Van der Ende, J. (1995). The eight-year stability of problem behavior in an epidemiologic sample. *Pediatric Research*, *38*, 612–617.

Woody-Ramsey, J., & Miller, P.H. (1988). The facilitation of selective attention in preschoolers. *Child Development*, *59*, 1497–1503.

Wothke, W. (2000). Longitudinal and multigroup modeling with missing data. In T.D. Little, K.U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data; Practical issues, applied approaches and specific examples*. Mahwah, NJ: Lawrence Erlbaum Associates.

Zahn-Waxler, C., Schmitz, S., Fulker, D., Robinson, J., & Emde, R. (1996). Behavior problems in 5-year-old monozygotic and dizygotic twins: Genetic and environmental influences, patterns of regulation, and internalizing of control. *Developmental Psychopathology*, *8*, 103–122.

Manuscript accepted 30 May 2003

Appendix I

Description of translated items for problem scales 'overactive behavior' and 'attention problems'

CBCL/2–3 for 3-year-olds Overactive behavior [OA]		CBCL/4–18 for 7-, 10-, and 12-year-olds Attention problems [AP]	
Item no.	Item description	Item no.	Item description
5.	Can't concentrate, can't pay attention for long	1.	Acts too young for his/her age
6.	Can't sit still, restless	8.	Can't concentrate, can't pay attention for long
11.	Constantly seeks help	10.	Can't sit still, restless, or overactive
59.	Quickly shifts from one activity to another	13.	Confused or seems to be in a fog
62.	Refuses to play active games	17.	Daydreams or gets lost in his/her thoughts
		41.	Impulsive or acts without thinking
		45.	Nervous or tense
		46	Nervous movements or twitching
		61.	Poor school work
		62.	Clumsy or poorly coordinated
		80.	Stares blankly